SEQUENCE LISTING

<110> Commissariat à l'Energie Atomique
Centre National de la Recherche Scientifique
GONDRY Muriel
GENET Roger
LAUTRU Sylvie
PERNODET Jean-Luc

<120> Polynucleotides and polypeptides coded by said polynucleotides involved in the synthesis of diketopiperazine derivatives

<130> CGA263/83FR

<140>
<141>

<160> 23

<170> PatentIn Ver. 2.1

<210> 1
<211> 657
<212> DNA
<213> Streptomyces noursei

<400> 1
gtgaggcgcc acccatcgca ttcgccgtac cgcggcgggt gtgaggtgcg cccaaaaaga 60
aggggattga tgttagctca cagttcatct gaatcgccgc cggaatcctt gccggacgcg 120
tggacggtcc tcaaaacccg taccgccgtc cgcaattacg cgaaagagcc ggtcgacgac 180
gcgctgatcg agcagctgtt ggaggccatg ctcgccgcgc cgaccgcctc caaccggcag 240
gcgtggtcgt tcatggtggt gcgcaggccc gccgcggtcc gccggctgcg cgcgttctcg 300
cccggggtgc tgggaacccc cgccttcttc gtcgtggcct gcgtcgaccg cagtctgacc 360
gacaacctct ccccgaagct ctcgcagaag atctacgaca ccagcaagct ctgtgtcgcc 420
atggcggtgg agaacctgct gctcgcggcg cacgcggccg gcctgggcgg atgcccggtg 480
ggcagcttca ggtccgacat cgtcaccagc atgctcggta tcccggaaca catcgagccg 540
atgctcgtgg tcccgatcgg ccgtcccgcg acagccctcg tcccctccca gcgccgcgcc 600
aagaatgagg tcgtcaacta tgaatcctgg ggaaaccgtg ctgccgcccc aactgcg 657

<210> 2
<211> 318
<212> DNA
<213> Streptomyces noursei

<400> 2
atgaatcctg gggaaaccgt gctgccgccc caactgcgtg aggagatcgc gctcctcgcc 60
gtctatctgc tcagcagcgg ccgcggactc ctggaggagc cggccgacta cggaatttac 120
cgctgtaccg acggggcccg tcgggcgctc caactcctcg acgaacacgg cgggagcacg 180
gcacggctga ccgccgtccg cgagcgtctc gacgaggtca tgttcgcgcc gatgggcgag 240
gaccgggaca tgggcgcgat tctggacgac ctgtgtcgcc aaatggcaga cgctcttccg 300
gaaattgaaa ccccctga 318

<210> 3
<211> 720
<212> DNA
<213> Streptomyces noursei

<400> 3
atgcttgcag gcttagttcc cgcgccggac cacggaatgc gggaagaaat acttggcgac 60
cgcagccgat tgatccggca acgcggtgag cacgccctca tcggaatcag tgcgggcaac 120
agttatttca gccagaagaa caccgtcatg ctgctgcaat gggccgggca gcgtttcgag 180
cgcaccgatg tcgtctatgt cgacacccac atcgacgaga tgctgatcgc cgacggccgc 240
agcgcgcagg aggccgagcg gtcggtcaaa cgcacgctca aggatctgcg gcgcagactc 300
cggcgctcgc tggagagcgt gggcgaccac gccgagcggt tccgtgtccg gtccctgtcc 360
gagctccagg agacccctga gtaccgggcc gtacgcgagc gcaccgaccg ggccttcgag 420

gaggacgccg aattcgccac cgcctgcgag gacatggtgc gggccgtggt gatgaaccgg 480
cccggtgacg gcgtcggcat ctccgcggaa cacctgcggg ccggtctgaa ctacgtgctg 540
gccgaggccc cgctcttcgc ggactcgccc ggagtcttct ccgtcccctc ctcggtgctc 600
tgctaccaca tcgacacccc gatcacggcg ttcctgtccc ggcgcgagac cggtttccgg 660
gcggccgagg gacaggcgta cgtcgtcgtc aggccccagg agctggccga cgcggcctag 720

<210> 4
<211> 834
<212> DNA
<213> Streptomyces noursei

<400> 4
atgtcatggg gaggacagga cacttgctca tggtgcggaa cggggcccct cggcgaagct 60
gaagacgtag gaagacagca cacgtcgcac gccgggggac ccgtcatgac tcaagccgcc 120
accgtcaccg ccaccacgag ccagggcagg gcactcctgc ggagcctgac gccgctgttc 180
gtggacgccg cgatcccgct cggctcgtac ttcctcctcg ccgagggctt cggcatgagc 240
acggtcgccg cgctggcctg gagcagcgtg gtcccggcgc tgcgcacgat ctggggcctg 300
gtccgggagc ggacggtcaa cggcctcgcg ctgctgatcc tcgtcgtcaa cgtggtgggg 360
ctggcgacga gcaccctgac cggcgatgcc cggctgatga tggccaagga cagcggcgtc 420
agcagcgtcg tcgggatcgc gatcctgctc tcggtgcgcg gccggcgccc gctgatgacc 480
gccggactcc ggccctgggt gaccaaggga agcccggagg ggaacgccgc atgggaccgg 540
ctgtgggcgc gcagcgcgcg gttccggcaa ctggagcggc gattctcgac ggtctggggg 600
agcgccctgc tgatcgagtg cgtggtcaag gtcgtcggtg cgtacgtcct gccggtgcac 660
accatggtgt ggctgggcac ggtgctgacg gtggtggcga tcctgctggc catggtggtc 720
gcgggcggcg gcagcgccga gccgatggag cggatggtca aggccgaggt cggggccgcc 780
ggcgaggccg ccacggcggg gaacgccgag ccggcgccgg ccgccgcggc ctga 834

<210> 5
<211> 3839
<212> DNA
<213> Streptomyces noursei

<400> 5
ggatccgtcc cgacgggcgg gaaccggtga ggcgccaccc atcgcattcg ccgtaccgcg 60
gcgggtgtga ggtgcgccca aaaagaaggg gattgatgtt agctcacagt tcatctgaat 120
cgccgccgga atccttgccg gacgcgtgga cggtcctcaa aacccgtacc gccgtccgca 180
attacgcgaa agagccggtc gacgacgcgc tgatcgagca gctgttggag gccatgctcg 240
ccgcgccgac cgcctccaac cggcaggcgt ggtcgttcat ggtggtgcgc aggcccgccg 300
cggtccgccg gctgcgcgcg ttctcgcccg gggtgctggg aacccccgcc ttcttcgtcg 360
tggcctgcgt cgaccgcagt ctgaccgaca acctctcccc gaagctctcg cagaagatct 420
acgacaccag caagctctgt gtcgccatgg cggtggagaa cctgctgctc gcggcgcacg 480
cggccggcct gggcggatgc ccggtgggca gcttcaggtc cgacatcgtc accagcatgc 540
tcggtatccc ggaacacatc gagccgatgc tcgtggtccc gatcggccgt cccgcgacag 600
ccctcgtccc ctcccagcgc cgcgccaaga atgaggtcgt caactatgaa tcctggggaa 660
accgtgctgc cgccccaact gcgtgaggag atcgcgctcc tcgccgtcta tctgctcagc 720
agcggccgcg gactcctgga ggagccggcc gactacggaa tttaccgctg taccgacggg 780
gcccgtcggg cgctccaact cctcgacgaa cacggcggga gcacggcacg gctgaccgcc 840
gtccgcgagc gtctcgacga ggtcatgttc gcgccgatgg gcgaggaccg ggacatgggc 900
gcgattctgg acgacctgtg tcgccaaatg gcagacgctc ttccggaaat tgaaaccccc 960
tgacggctgt ccggggcaac cccaaaagga cttcttagca tgcttgcagg cttagttccc 1020
gcgccggacc acggaatgcg ggaagaaata cttggcgacc gcagccgatt gatccggcaa 1080
cgcggtgagc acgccctcat cggaatcagt gcgggcaaca gttatttcag ccagaagaac 1140
accgtcatgc tgctgcaatg ggccgggcag cgtttcgagc gcaccgatgt cgtctatgtc 1200
gacacccaca tcgacgagat gctgatcgcc gacggccgca gcgcgcagga ggccgagcgg 1260
tcggtcaaac gcacgctcaa ggatctgcgg cgcagactcc ggcgctcgct ggagagcgtg 1320
ggcgaccacg ccgagcggtt ccgtgtccgg tccctgtccg agctccagga gacccctgag 1380
taccgggccg tacgcgagcg caccgaccgg gccttcgagg aggacgccga attcgccacc 1440
gcctgcgagg acatggtgcg ggccgtggtg atgaaccggc ccggtgacgg cgtcggcatc 1500
tccgcggaac acctgcgggc cggtctgaac tacgtgctgg ccgaggcccc gctcttcgcg 1560
gactcgcccg gagtcttctc cgtcccctcc tcggtgctct gctaccacat cgacaccccg 1620
atcacggcgt tcctgtcccg gcgcgagacc ggtttccggg cggccgaggg acaggcgtac 1680
gtcgtcgtca ggccccagga gctggccgac gcggcctagt tgggggcgtc cgcgggcgga 1740
cctgcctccc cacccgctcc cggtgccggc gccgggcatg acaaatgtca tggggaggac 1800
aggacacttg ctcatggtgc ggaacggggc ccctcggcga agctgaagac gtaggaagac 1860
agcacacgtc gcacgccggg ggacccgtca tgactcaagc cgccaccgtc accgccacca 1920
cgagccaggg cagggcactc ctgcggagcc tgacgccgct gttcgtggac gccgcgatcc 1980

```
cgctcggctc gtacttcctc ctcgccgagg gcttcggcat gagcacggtc gccgcgctgg 2040
cctggagcag cgtggtcccg gcgctgcgca cgatctgggg cctggtccgg gagcggacgg 2100
tcaacggcct cgcgctgctg atcctcgtcg tcaacgtggt ggggctggcg acgagcaccc 2160
tgaccggcga tgcccggctg atgatggcca aggacagcgg cgtcagcagc gtcgtcggga 2220
tcgcgatcct gctctcggtg cgcggccggc gcccgctgat gaccgccgga ctccggccct 2280
gggtgaccaa gggaagcccg gaggggaacg ccgcatggga ccggctgtgg gcgcgcagcg 2340
cgcggttccg gcaactggag cggcgattct cgacggtctg ggggagcgcc ctgctgatcg 2400
agtgcgtggt caaggtcgtc ggtgcgtacg tcctgccggt gcacaccatg gtgtggctgg 2460
gcacggtgct gacggtggtg gcgatcctgc tggccatggt ggtcgcgggc ggcggcagcg 2520
ccgagccgat ggagcggatg gtcaaggccg aggtcggggc cgccggcgag gccgccacgg 2580
cggggaacgc cgagccggcg ccggccgccg cggcctgaga ccgcgcggcg ggggagttgg 2640
ggaaatcgcg tacaggattg gcgcgtcgag cacccccgcc ctcgataggg cgggcccccg 2700
gcgcatatgg tcggcgatgc gacgggacat cggagccccg cgtcgacggt tcaacggcga 2760
tccggacggc acgcggcttt cgtcggccac gaagggaacg gaagtcatgt cgactgttca 2820
cactggggtc acgcagagcg gtctcaccgc cgagctggcc tccctgcacg ccgagctcgt 2880
ccgtcggaat cccggtgaag cggagttcca ccaggcggcc ctggaggtcc tcgaaacgct 2940
ggcaccggtg ctcaccgccc ggccggagtt cgccgacgcc aaggtcctgg agcggatcgt 3000
cgagccggag cggcagatca tgttccgcgt gccctggcag gacgactccg gcacgatccg 3060
ggtcaaccgc ggcttccggg tggagttcaa cagcgcgctc ggcccctaca agggcggcct 3120
gcggttccac gcgtccgtca acctcggcat cgtgaagttc ctcggcttcg agcagatctt 3180
caagaacgcc ctgaccgggc tgaacatcgg cggcggcaag ggcggcagcg acttcgaccc 3240
gcacggcagg tcggacgccg aggtgatgcg cttctgccag tccttcatga ccgagctgca 3300
ccgtcacctg ggcgagcaca ccgacgtgcc ggctggcgac atcggcgtcg gcggccggga 3360
gatcggctac ctcttcggcc agtaccggcg gatcaccaac cgctgggagg ccggcgtcct 3420
gaccggcaag ggcctggcgt ggggcggctc caaggcccgt acggaggcca ccggttacgg 3480
caatgtgctg ttcaccgagg agatgctcaa gcagcgcggc gaggagctgg acggccagca 3540
ggtggtggtc tccgggtccg gcaacgtcgc catctacacc atcgagaagg cccaggcgct 3600
cggcgccaac gtcctgaccg tctcggactc cggcggctac gtcgtcgacg agaagggcat 3660
cgacctggcg ctgctcaagc aggtcaagga ggtcgagcgc ggccgggtcg gcgactacgc 3720
ccagcggcgc ggcagttcgg cgaagtacgt cgccggcggg agcgtgtggg acgtcgcctg 3780
tgacgtggcg ctgccgtcgg ccacccagaa cgagctcgac gcggacgccg cccggatcc  3839
```

<210> 6
<211> 219
<212> PRT
<213> Streptomyces noursei

<400> 6

```
Met Arg Arg His Pro Ser His Ser Pro Tyr Arg Gly Gly Cys Glu Val
  1               5                  10                  15

Arg Pro Lys Arg Arg Gly Leu Met Leu Ala His Ser Ser Ser Glu Ser
            20                  25                  30

Pro Pro Glu Ser Leu Pro Asp Ala Trp Thr Val Leu Lys Thr Arg Thr
        35                  40                  45

Ala Val Arg Asn Tyr Ala Lys Glu Pro Val Asp Asp Ala Leu Ile Glu
    50                  55                  60

Gln Leu Leu Glu Ala Met Leu Ala Ala Pro Thr Ala Ser Asn Arg Gln
65                  70                  75                  80

Ala Trp Ser Phe Met Val Val Arg Arg Pro Ala Ala Val Arg Arg Leu
                85                  90                  95

Arg Ala Phe Ser Pro Gly Val Leu Gly Thr Pro Ala Phe Phe Val Val
            100                 105                 110

Ala Cys Val Asp Arg Ser Leu Thr Asp Asn Leu Ser Pro Lys Leu Ser
        115                 120                 125

Gln Lys Ile Tyr Asp Thr Ser Lys Leu Cys Val Ala Met Ala Val Glu
    130                 135                 140

Asn Leu Leu Leu Ala Ala His Ala Ala Gly Leu Gly Gly Cys Pro Val
145                 150                 155                 160
```

```
Gly Ser Phe Arg Ser Asp Ile Val Thr Ser Met Leu Gly Ile Pro Glu
                165                 170                 175

His Ile Glu Pro Met Leu Val Val Pro Ile Gly Arg Pro Ala Thr Ala
            180                 185                 190

Leu Val Pro Ser Gln Arg Arg Ala Lys Asn Glu Val Val Asn Tyr Glu
        195                 200                 205

Ser Trp Gly Asn Arg Ala Ala Ala Pro Thr Ala
    210                 215
```

<210> 7
<211> 104
<212> PRT
<213> Streptomyces noursei

<400> 7

```
Asn Pro Gly Glu Thr Val Leu Pro Pro Gln Leu Arg Glu Glu Ile Ala
  1               5                  10                  15

Leu Leu Ala Val Tyr Leu Leu Ser Ser Gly Arg Gly Leu Leu Glu Glu
            20                  25                  30

Pro Ala Asp Tyr Gly Ile Tyr Arg Cys Thr Asp Gly Ala Arg Arg Ala
        35                  40                  45

Leu Gln Leu Leu Asp Glu His Gly Gly Ser Thr Ala Arg Leu Thr Ala
    50                  55                  60

Val Arg Glu Arg Leu Asp Glu Val Met Phe Ala Pro Met Gly Glu Asp
65                  70                  75                  80

Arg Asp Met Gly Ala Ile Leu Asp Asp Leu Cys Arg Gln Met Ala Asp
                85                  90                  95

Ala Leu Pro Glu Ile Glu Thr Pro
            100
```

<210> 8
<211> 99
<212> PRT
<213> Streptomyces noursei

<400> 8

```
Met Leu Pro Pro Gln Leu Arg Glu Glu Ile Ala Leu Leu Ala Val Tyr
  1               5                  10                  15

Leu Leu Ser Ser Gly Arg Gly Leu Leu Glu Glu Pro Ala Asp Tyr Gly
            20                  25                  30

Ile Tyr Arg Cys Thr Asp Gly Ala Arg Arg Ala Leu Gln Leu Leu Asp
        35                  40                  45

Glu His Gly Gly Ser Thr Ala Arg Leu Thr Ala Val Arg Glu Arg Leu
    50                  55                  60

Asp Glu Val Met Phe Ala Pro Met Gly Glu Asp Arg Asp Met Gly Ala
65                  70                  75                  80

Ile Leu Asp Asp Leu Cys Arg Gln Met Ala Asp Ala Leu Pro Glu Ile
                85                  90                  95

Glu Thr Pro
```

<210> 9
<211> 239
<212> PRT
<213> Streptomyces noursei

<400> 9
Met Leu Ala Gly Leu Val Pro Ala Pro Asp His Gly Met Arg Glu Glu
1 5 10 15

Ile Leu Gly Asp Arg Ser Arg Leu Ile Arg Gln Arg Gly Glu His Ala
20 25 30

Leu Ile Gly Ile Ser Ala Gly Asn Ser Tyr Phe Ser Gln Lys Asn Thr
35 40 45

Val Met Leu Leu Gln Trp Ala Gly Gln Arg Phe Glu Arg Thr Asp Val
50 55 60

Val Tyr Val Asp Thr His Ile Asp Glu Met Leu Ile Ala Asp Gly Arg
65 70 75 80

Ser Ala Gln Glu Ala Glu Arg Ser Val Lys Arg Thr Leu Lys Asp Leu
85 90 95

Arg Arg Arg Leu Arg Arg Ser Leu Glu Ser Val Gly Asp His Ala Glu
100 105 110

Arg Phe Arg Val Arg Ser Leu Ser Glu Leu Gln Glu Thr Pro Glu Tyr
115 120 125

Arg Ala Val Arg Glu Arg Thr Asp Arg Ala Phe Glu Glu Asp Ala Glu
130 135 140

Phe Ala Thr Ala Cys Glu Asp Met Val Arg Ala Val Val Met Asn Arg
145 150 155 160

Pro Gly Asp Gly Val Gly Ile Ser Ala Glu His Leu Arg Ala Gly Leu
165 170 175

Asn Tyr Val Leu Ala Glu Ala Pro Leu Phe Ala Asp Ser Pro Gly Val
180 185 190

Phe Ser Val Pro Ser Ser Val Leu Cys Tyr His Ile Asp Thr Pro Ile
195 200 205

Thr Ala Phe Leu Ser Arg Arg Glu Thr Gly Phe Arg Ala Ala Glu Gly
210 215 220

Gln Ala Tyr Val Val Val Arg Pro Gln Glu Leu Ala Asp Ala Ala
225 230 235

<210> 10
<211> 277
<212> PRT
<213> Streptomyces noursei

<400> 10
Met Ser Trp Gly Gly Gln Asp Thr Cys Ser Trp Cys Gly Thr Gly Pro
1 5 10 15

Leu Gly Glu Ala Glu Asp Val Gly Arg Gln His Thr Ser His Ala Gly
20 25 30

Gly Pro Val Met Thr Gln Ala Ala Thr Val Thr Ala Thr Thr Ser Gln
35 40 45

```
Gly Arg Ala Leu Leu Arg Ser Leu Thr Pro Leu Phe Val Asp Ala Ala
     50                  55                  60

Ile Pro Leu Gly Ser Tyr Phe Leu Leu Ala Glu Gly Phe Gly Met Ser
 65                  70                  75                  80

Thr Val Ala Ala Leu Ala Trp Ser Ser Val Val Pro Ala Leu Arg Thr
                85                  90                  95

Ile Trp Gly Leu Val Arg Glu Arg Thr Val Asn Gly Leu Ala Leu Leu
            100                 105                 110

Ile Leu Val Val Asn Val Val Gly Leu Ala Thr Ser Thr Leu Thr Gly
        115                 120                 125

Asp Ala Arg Leu Met Met Ala Lys Asp Ser Gly Val Ser Ser Val Val
    130                 135                 140

Gly Ile Ala Ile Leu Leu Ser Val Arg Gly Arg Arg Pro Leu Met Thr
145                 150                 155                 160

Ala Gly Leu Arg Pro Trp Val Thr Lys Gly Ser Pro Glu Gly Asn Ala
                165                 170                 175

Ala Trp Asp Arg Leu Trp Ala Arg Ser Ala Arg Phe Arg Gln Leu Glu
            180                 185                 190

Arg Arg Phe Ser Thr Val Trp Gly Ser Ala Leu Leu Ile Glu Cys Val
        195                 200                 205

Val Lys Val Val Gly Ala Tyr Val Leu Pro Val His Thr Met Val Trp
    210                 215                 220

Leu Gly Thr Val Leu Thr Val Val Ala Ile Leu Leu Ala Met Val Val
225                 230                 235                 240

Ala Gly Gly Gly Ser Ala Glu Pro Met Glu Arg Met Val Lys Ala Glu
                245                 250                 255

Val Gly Ala Ala Gly Glu Ala Ala Thr Ala Gly Asn Ala Glu Pro Ala
            260                 265                 270

Pro Ala Ala Ala Ala
        275


<210> 11
<211> 20
<212> PRT
<213> Streptomyces noursei

<400> 11
Glu Pro Val Asp Asp Ala Leu Ile Glu Gln Leu Leu Glu Ala Met Leu
  1               5                  10                  15

Ala Ala Pro Thr
             20


<210> 12
<211> 12
<212> PRT
<213> Streptomyces noursei

<400> 12
Asn Glu Val Val Asn Tyr Glu Xaa Trp Gly Asn Arg
  1               5                  10
```

<210> 13
<211> 9
<212> PRT
<213> Streptomyces noursei

<400> 13
Gln Ala Xaa Ser Phe Met Val Val Arg
1 5

<210> 14
<211> 17
<212> DNA
<213> Artifical sequence

<220>
<223> Description of the artificial sequence:primer

<400> 14
garccsgtsg acgacgc 17

<210> 15
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 15
gcgtcgtcsa csggytc 17

<210> 16
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 16
aacgargtsg tsaactacga 20

<210> 17
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 17
tcgtagttsa csacytcgtt 20

<210> 18
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 18
caggcstggw ssttcatggt 20

<210> 19
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 19
accatgaass wccasgcctg 20

<210> 20
<211> 47
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 20
cggctgcagg agaagggagc ggacatatgc ttgcaggctt agttccc 47

<210> 21
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 21
cggtcccgtg gatccaagct tctaggccgc gtcggccagc tc 42

<210> 22
<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 22
gagcgggatc ctgcagtgtc atggggagga caggac 36

<210> 23
<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:primer

<400> 23
cgatcacgtg gatccaagct tgccaatcct gtacgcgatt t 41